

Building Data Analytics Solutions Using Amazon Redshift

AWS 課堂培訓

課程說明

在本課程中，您將使用 Amazon Redshift (一種雲端資料倉儲服務) 建置資料分析解決方案。本課程著重於分析管道的資料收集、擷取、編目、儲存和處理元件。您將學習如何將 Amazon Redshift 與資料湖整合，以支援分析和機器學習工作負載。您還將學習如何將安全性、效能和成本管理最佳實務套用至 Amazon Redshift 的操作。

- 課程級別：中級
- 持續時間：1 天

活動

本課程包括演示文稿，互動演示，實作實驗室，討論和課堂練習。

課程目標

在本課程中，您將學習：

- 比較資料倉儲、資料湖和現代資料架構的功能與優勢
- 設計和實作資料倉儲分析解決方案
- 識別並應用適當的技術，包括壓縮，以優化數據存儲
- 選取並部署適當的選項，以擷取、轉換和儲存資料
- 針對特定商業使用案例，選擇適當的執行個體和節點類型、叢集、自動擴展和網路拓撲
- 瞭解資料儲存和處理如何影響獲得可行商業洞察所需的分析和視覺化機制
- 保護靜態和傳輸中的資料
- 監控分析工作負載以識別並修復問題
- 套用成本管理最佳實務

目標受眾

本課程適用於：

Building Data Analytics Solutions Using Amazon Redshift

AWS 課堂培訓

- 資料倉儲工程師
- 資料平台工程師
- 建置和管理資料分析管道的架構師和營運商

前提

擁有至少一年管理資料倉儲經驗的學生將受惠於本課程。我們建議參加本課程的參加者：

- 已完成 [AWS 技術基礎要點](#)或在 [AWS 上架構](#)
- [在 AWS 上完成的建置資料湖](#)

課程大綱

單元 A：數據分析和數據管道概述

- 資料分析使用案例
- 使用資料管道進行分析

單元 1：在資料分析管道中使用 Amazon Redshift

- 使用 Amazon Redshift 進行資料倉儲的原因
- Amazon Redshift 概述
- Amazon Redshift 架構
- 交互式演示 1：遊覽 Amazon Redshift 控制台
- Amazon Redshift 功能
- 實驗室 1：在 Amazon Redshift 叢集中載入和查詢資料

Building Data Analytics Solutions Using Amazon Redshift

AWS 課堂培訓

單元 2：擷取與儲存

- 攝入
- 互動式示範 2：使用具有資料 API 的 Jupyter notebook 連接您的 Amazon Redshift 叢集
- 數據分發和存儲
- 互動式示範 3：使用 SUPER 資料類型分析半結構化資料
- 在 Amazon Redshift 中查詢數據
- 練習實驗室 2：使用 Amazon Redshift Spectrum 進行資料分析

單元 3：處理和最佳化資料

- 資料轉換
- 進階查詢
- 練習實驗室 3：在 Amazon Redshift 中進行資料轉換和查詢
- 資源管理
- 互動式示範 4：在 Amazon Redshift 上套用混合式工作負載管理
- 自動化與最佳化
- 互動式示範 5：Amazon Redshift 叢集從 dc2.large 調整為 ra3.xlplus 叢集的大小

單元 4：安全性與監控

- 保護 Amazon Redshift 叢集
- 監控和疑難排解 Amazon Redshift 叢集

單元 5：設計資料倉儲分析解決方案

- 資料倉儲使用案例檢閱
- 活動：設計資料倉儲分析工作流程

單元 B：在 AWS 上開發現代資料架構

- 現代化資料架構